Spatio-temporal integration properties of the human visual system

Grillini, Alessandro

DOI:
10.33612/diss.136424282

Chapter 3

Classification of visual field defects based on the spatio-temporal properties of eye-movements

Published as:
Grillini, A., Ombelet, D., Soans, R. S. and Cornelissen, F. W. (2018). Towards Using the Spatio-temporal Properties of Eye Movements to Classify Visual Field Defects. *In ETRA '18: 2018 Symposium on Eye Tracking Research and Applications, June 14–17, 2018, Warsaw, Poland. ACM, New York, NY, USA, Article 4, 5 pages.* doi:/10.1145/3204493.3204590

Perimetry—assessment of visual field defects (VFD)—requires patients to be able to maintain a prolonged stable fixation, as well as to provide feedback through motor response. These aspects limit the testable population and often lead to inaccurate results. We hypothesized that different VFD would alter the eye-movements in systematic ways, thus making it possible to infer the presence of VFD by quantifying the spatio-temporal properties of eye movements. We developed a tracking test to record participants' eye movements while we simulated different gaze-contingent VFD. We tested 50 visually healthy participants and simulated three common scotomas: peripheral loss, central loss and hemifield loss. We quantified spatio-temporal features using cross-correlogram analysis, then applied cross-validation to train a decision tree algorithm to classify the conditions. Our test is faster and more comfortable than standard perimetry and can achieve a classifying accuracy of 90% (True Positive Rate = 98%) with data acquired in less than 2 minutes.

3.1 Introduction

Detecting the presence of visual field defects (VFD) at an early stage is a critical aspect of diagnosing many ophthalmic and neurological disorders (i.e. glaucoma, macular degeneration, post-stroke visual loss). There are several ways to assess the presence of VFD. Two of the most commonly used methods are Standard Automated Perimetry (SAP) and Frequency Doubling Technology (FDT) perimetry. However, there are a number of challenges when it comes to visual field testing. Firstly, most test versions of SAP require concentration for a prolonged period of time ($\sim$ 10 min)—this can render the task tedious for some participants and impossible for others. FDT instead takes approx. 2 min but is relatively insensitive to minor defects. Secondly, in both SAP and FDT, the participant is required to maintain stable fixation. Failing to do so results in fixation loss and—consequently—an extended measurement of time or less reliable results. Thirdly, all current perimetric techniques require the patients to provide their feedback through motor response. While being no more than a nuisance to many participants, these limitations prevent the clinical assessment of some groups of patients. In particular, elderly people and children are affected by these limitations.

Modern perimetry has its roots in psychophysical methodology. The participant has to detect a stimulus and answer the simple question: "Did you see it?" However, the continuous need to question the participant repeatedly is one of the reasons that extend the assessment period. A solution could be continuously interrogating the visual system's ability to detect the presence of the stimulus. If participants are able to detect a stimulus, they also have information about its approximate position. While if the stimulus is less visible or invisible, its positional information will also deteriorate or be absent. This suggests that an alternative means of assessing the visual system lies in the continuous tracking of a stimulus position. Indeed, previous studies showed that the spatial sensitivity measured by continuous tracking (performed with a computer mouse) correlates well with the one measured by means of traditional psychophysics[26].

However, the human visual system has a "built-in tracking device"—i.e. the eyes—of which the performance is very directly linked to the quality of visual processing. For this reason, we studied the possibility of using eye-movement properties assessed by means of a continuous tracking task to screen for the presence of VFD. We hypothesize that different VFD will alter the spatio-temporal eye movements in a specific and measurable way. We expect that the presence of a VFD will affect one's ability to track a stimulus. First, if the sensitivity of the central visual field or the entire visual system is reduced, we can expect prolonged processing time and therefore delayed

tracking performance. Second, if the tracking stimulus completely disappears because of a scotoma, it will take time to find it again and pursue tracking, delaying the performance and increasing the spatial error. Given these predicted alterations, we expect that it should also be possible to identify the underlying defect based on the spatio-temporal characteristics of the eye movements by training a machine-learning classifier.

3.2 Methods

3.2.1 Participants and Ethical Clearance

We tested 50 healthy participants and one glaucoma patient with peripheral visual field loss, previously assessed with SAP. All had normal or corrected-to-normal visual acuity, verified prior to data collection with "FrACT"[28]. The study followed the tenets of the Declaration of Helsinki. The ethics board of the Psychology Department of the University of Groningen approved the study protocol. All participants provided written informed consent before participation.

3.2.2 Procedures

Participants performed a visual tracking task while being subjected to different simulations of VFD. No simulation was applied to the glaucoma patient. The conditions were sorted with stratified randomization. The gaze position was constantly monitored and recorded.

Apparatus

The experiment was designed and conducted with custom-made scripts in MATLAB using Psychtoolbox[29] and Eyelink Toolbox[30]. The data was acquired with an eye-tracker Eyelink 1000 (SR-Research, Kanata, Ontario, Canada) with a sampling frequency of 1 KHz, downsampled to 240 Hz to match the refresh rate of the stimulus display monitor Zowie xl2540 (BenQ, Taipei, Taiwan).

Tracking Stimulus

The stimulus comprises of a Gaussian luminance blob (FWHM = 0.5 degrees of visual angle) moving in a 2D random-walk path. In each trial, the stimulus could be presented either at a high (50%) or low (5%) contrast. The random-walk path could be smooth (continuous) or saccadic (random displacement applied every 2 sec). For the smooth path, the instantaneous speeds for every frame of the stimulus presentation were randomly drawn from a standard normal distribution, then multiplied by two different gains for the horizontal and vertical components (drawn independently). Both the resulting distributions are zero-mean and have standard deviations of $\pm$64.45 deg/sec (horizontal) and $\pm$32.23 deg/sec (vertical). For the saccadic path, the random displacement is achieved by juxtaposing epochs of 2 seconds each taken from the original 6 smooth trajectories, with the only constraint that the stimulus must never land outside the screen boundaries.

Gaze-contingent simulated VFD

The simulations were obtained by superimposing in real-time (240 Hz, 4 ms delay) a uniform grey area to the current position of the participant's gaze. The shape and size of the simulated VFD were modeled after the typical scotoma resulting from three common ophthalmologic disorders: age-related macular degeneration (central loss), late stage glaucoma (peripheral loss) and hemianopia (hemifield loss) (Figure 3.1). For each combination of conditions (pursuit $\times$ contrast $\times$ VFD), we performed 6 tracking trials of 20 sec each.

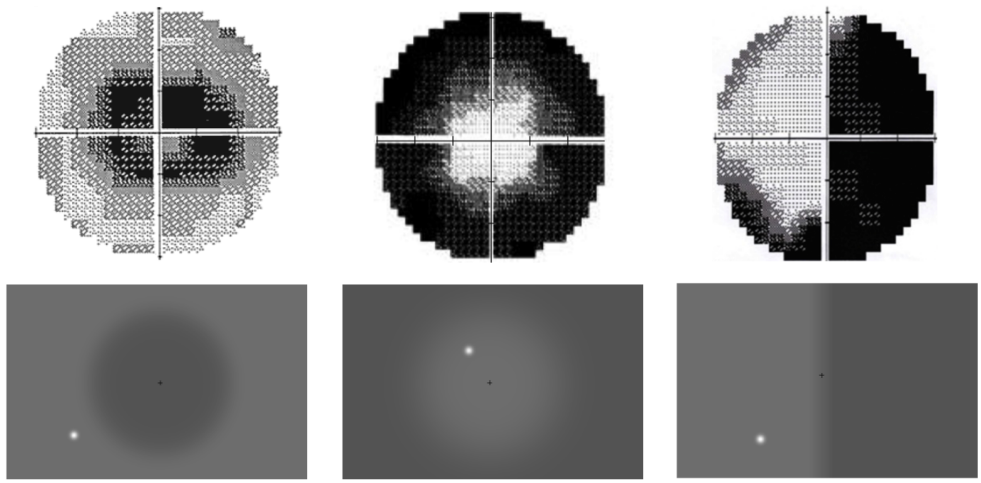

Figure 3.1: Simulated visual field defects.
Top row: examples of scotoma represented by visual field maps. From left to right: central loss, peripheral loss, hemifield loss. Bottom row: simulated gaze-contingent VFD (darker for visualization purposes only).

3.2.3 Spatio-temporal Features Extraction

Abbreviations:

- **CCG1** Cross-correlogram between the smooth velocity of the eye and stimulus;
- **CCG2** Cross-correlogram between saccadic velocity and positional error;
- **PDD** Probability density distribution of positional error.

The data was analyzed according to the Eye-Movement Correlogram technique described by Mulligan and colleagues, 2013[2]. The instantaneous velocities of stimulus and eye are thresholded to separate the saccadic component ($v(t) > 30$ deg/sec) from the smooth pursuit component, then cross-correlated independently and averaged across the 6 tracking trials per each condition. Finally, we computed the instantaneous positional error as the Euclidean Distance between the eye and stimulus positions.

The 10 spatio-temporal features that we used to train the classifier are: *amplitude*, μ, σ, and R^2 of the Gaussian fits to CCG1 and to PDD, respectively, and *peak* and *delay* of the CCG2. All features were determined separately for the horizontal and the vertical components of the eye-movements. Therefore, we obtained a dataset comprising 80 features (10 features $\times$ 2 axis $\times$ 2 contrast levels $\times$ 2 pursuit modalities) per 200 entries (50 participants $\times$ 4 conditions).

3.2.4 Features Classification

We used Decision Trees (DTs)[31] for discriminating between the various VFD and controls. Since DTs have an inherent feature selection capability[32], we used them also for reducing the dimensionality of our dataset. DTs split a set of labeled data at each node into smaller branches before eventually converging onto one of the possible decisions. In our context, the initial node is an 80-dimensional feature space that is constituted from the 80 features obtained from the CCG and PDD analysis. The feature space is then split into smaller sub-spaces at every subsequent node based on an "impurity criterion".

The decision criterion that we use is the "Gini's Diversity Index" (GDI) given as follows:

$$1 - \sum f^2(i) \tag{3.1}$$

Where *f(i)* is the fraction of the number of samples in the training set for class *i* that reaches a particular node (*i* being an index for control, peripheral loss, central loss, and hemifield loss). If a node has samples of only 1 class coming into it, then the GDI has a value of 0 and the node is termed as a pure node. Consequently, that node is assigned as a class, thereby completing the decision. On the other hand, if a node has a mixture of samples coming from various classes, then the GDI has some positive value and the node is split in such a way so as to minimize the GDI after the split. Finally, to estimate the classifier's performance, we used a 10-fold cross-validation scheme wherein the entire dataset is randomly partitioned into 10 subsets of equal size. Subsequently, 9 subsets constitute the training set, and the remaining 1 set is used as test set. This process is then repeated until every subset has been used once as a test set. The estimated total accuracy is the average of the accuracies measured after each repetition.

3.3 Results

3.3.1 Cross-Correlograms and Probability Density Distributions

Figure 3.2 shows an example of analyzed data for one participant in the smooth pursuit (panel A) and saccadic pursuit (panel B) tracking modalities, in the control and the three VFD simulation conditions, respectively. The top row shows CCG1 (stimulus $*$ eye velocities). The Gaussian fits for the simulated VFD (in blue, yellow and red) are

both delayed and shallower compared to those for the control condition (in green), indicating reduced performance in both pursuit modalities (A and B). The mid-row shows CCG2 (positional error $*$ saccadic velocity). The peak of the correlation at time point 0 is negative because the saccadic velocity is always of the opposite sign of the positional error: if the eye is overshooting to the right (positive positional error) correction requires making a saccade to the left and vice-versa. Also, in this case, the CCG gets shallower in the VFD conditions, indicating longer correction time. The bottom row shows the positional error (PDD). Again, the VFD leads to shallower Gaussian fits indicating performance deterioration. Noticeably, in the case of hemifield loss, we also observe a large shift of the mean error in the direction of the scotoma.

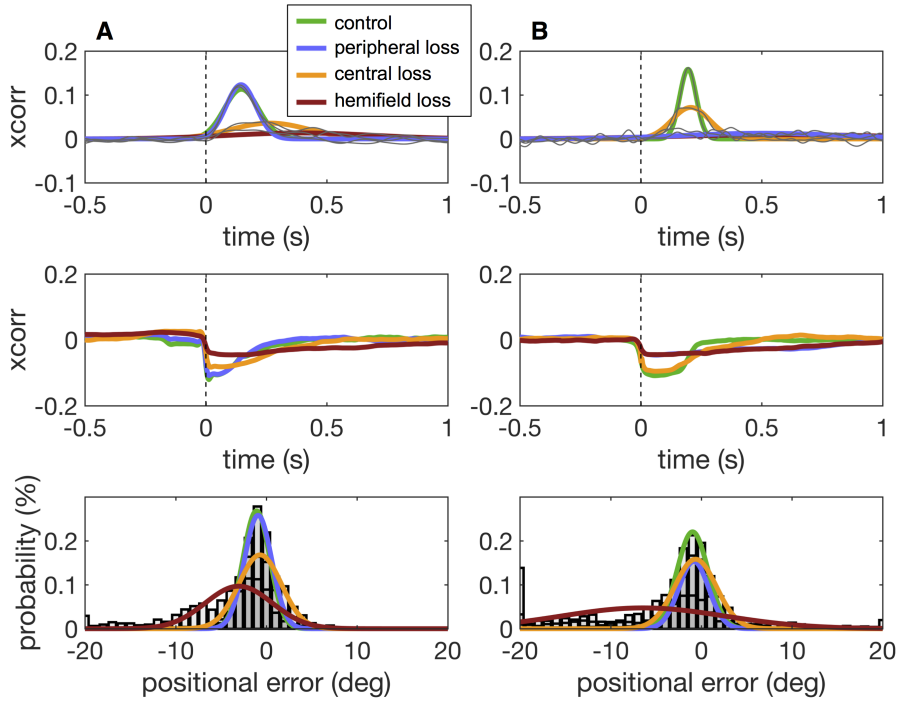

Figure 3.2: Example of single participant results.
A. Results for the smooth pursuit condition of a single participant. **B**. Results for the saccadic pursuit condition of a single participant. CCG1 (top) is the cross-correlation between the stimulus' and the eye's velocities, CCG2 (middle) is the cross-correlation between the positional error and the saccadic velocity, and PDD (bottom) is the probability distribution of positional errors. The spatiotemporal features are extracted from the resulting CCGs and PDD.

3.3.2 Features Selection and Classifier Performance

Figure 3.3 shows the DT resulting from the highest accuracy achieved (91.5%). The most noticeable thing is the large reduction in dimensionality: from the initial 80 features only 4 "survived". Moreover, we observed that amongst the different features, the vertical axis and the high contrast ones were always absent, thus not contributing to the classification. Therefore, these feature levels are omitted from the graphical representation shown in Figure 3.3.

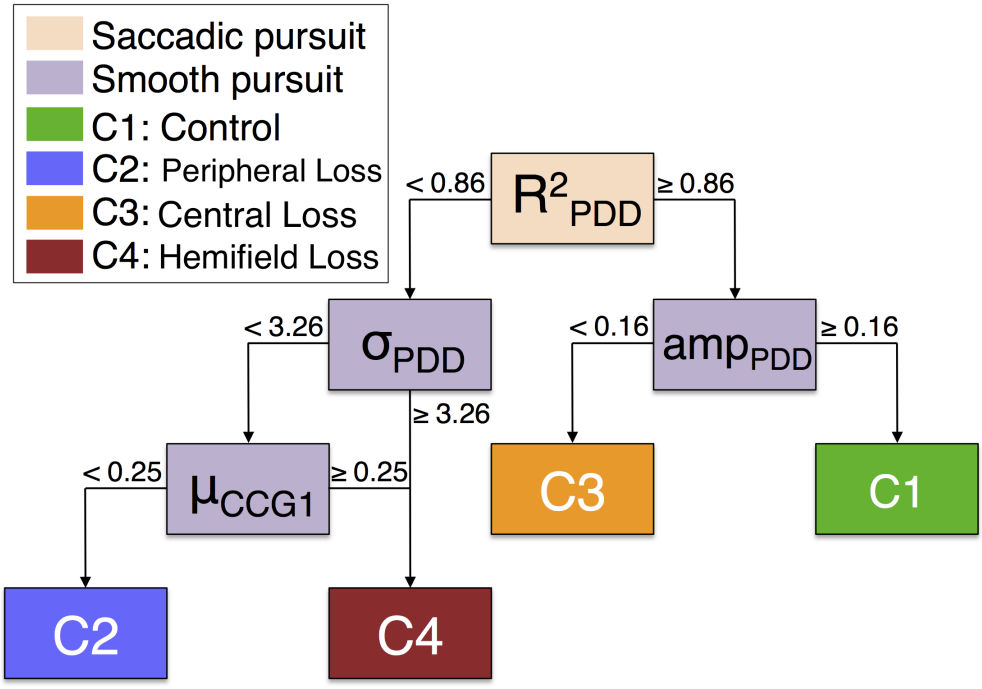

Figure 3.3: Decision Tree to classify visual field defects.
The features used in the classification are saccadic R2, standard deviation, and amplitude of smooth pursuit PDD, and the mean of CCG1. The DT finds the threshold values in these features that achieve the best separation between classes.

Figure 3.4-A shows the accuracy of the classifier as a function of the duration of the data collection. Even just 20 seconds of data collection is sufficient to yield an accuracy level above 75%. The accuracy stabilizes at around 90% for duration above 1 min. Finally, Figure 3.4-B shows the True Positive Rate (TPR; i.e., "when a participant really has a disorder, how often our test classifies it correctly?"), a measure that is of paramount importance in clinical screening tests. We found a very high TPR for controls (98%), and relatively lower TPRs for each VFD condition.

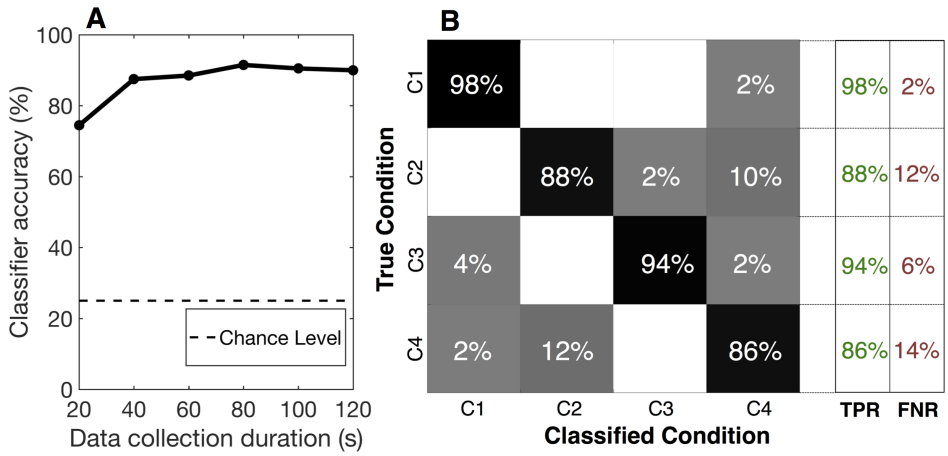

Figure 3.4: Performance of the Decision Tree classifier
A. Classifier accuracy. **B**. True Positive Rate. C1: control; C2: central loss; C3: peripheral loss; C4: hemifield loss.

3.3.3 Comparison between simulated and real peripheral loss

Figure 3.5 shows the comparison between the control group, simulated peripheral loss group, and a single patient previously diagnosed with peripheral visual field loss due to glaucoma. Only parameters derived from the saccadic pursuit are shown since for peripheral loss the smooth pursuit condition results are very similar to controls (see Figure 3.2, top row, the overlap between blue and green lines). The pattern obtained with the glaucoma patient closely resembles the simulated peripheral loss group pattern, rather than the control group one.

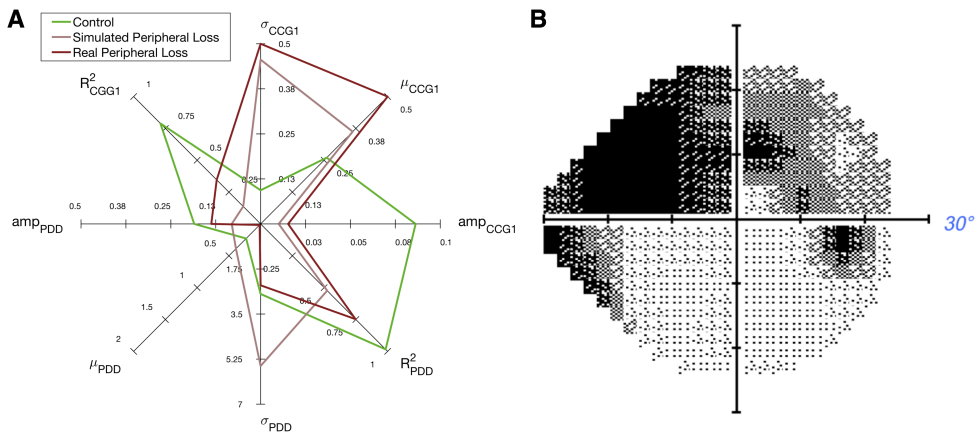

Figure 3.5: Glaucoma patient's results
A. Patterns obtained from spatio-temporal features. The patient pattern (in red) resembles the average pattern obtained from simulated VFD, while the control group is clearly separated. **B**. Corresponding visual field.

3.4 Discussion

Our main finding is that the spatio-temporal properties of eye movements change when a simulated VFD interferes with the tracking task. Further, these changes appear to be systematic and specific to each VFD. Moreover, we show that a classification approach based on an analysis of the spatio-temporal properties of eye movements is feasible.

3.4.1 Relevant and irrelevant screening conditions

Following the CCG and PDD analysis, we can observe that the differences between smooth pursuit and saccadic pursuit are more prominent in the peripheral loss condition and, partially, in the hemifield loss one. The most likely explanation is that during the saccadic task, the stimulus "jumping" into the scotomatous region made this task much more difficult. On the other hand, having their central vision spared, the participants were able to track the stimulus with ease in the smooth pursuit condition. This aspect needs to be taken into account if the goal is to develop a comprehensive screening test, and it is further confirmed by the DT results: both pursuit tasks contributed to correct classification, while other tested features turned out to be uninformative.

3.4.2 Clinical implications and limitations

A fast screening test that neither requires a manual motor response nor stable fixation potentially holds a major step forward from current SAP techniques. However, the present study was limited in various ways. First, we used simulations that are a simplified version of real VFD. Although our preliminary clinical results are encouraging, a thorough evaluation of patients with actual VFD, and a comparison to other perimetric techniques will be the next necessary step. Moreover, it has previously been shown that with either simulated or real defects, the eye movements will adapt to accommodate for various problems[33–35]. However, relying entirely on these accommodations in order to detect VFD is still a relatively unexplored field. A previous study managed to detect macular degeneration from the scan-path generated during natural viewing behavior[36]. Another limitation is the dependence on eye-tracker calibration, a procedure that can be difficult for a patient with severe central vision loss. Although the temporal features are not affected, since they are computed on relative displacements rather than absolute position, the spatial features are particularly sensitive to measurement bias.

The overall accuracy achieved by the classifier, tied to the very short amount of time needed, makes this potentially appealing to implement this as a clinical test. However, it is important to note that the number of participants was relatively small, compared to standard clinical and normative datasets. Furthermore, each participant served both as control and as "patient", hence removing a part of inter-subject variability typically present in the general population. More data and testing in an actual clinical setting would be required to generalize our results.

3.4.3 Conclusions

We have developed a new test for screening for visual field defects that is entirely based on evaluating eye movement behavior in response to an extremely simple visual task. This way, it improves on some of the most critical aspects of current perimetric techniques: it is faster and therefore hardly fatiguing, it is less attentionally taxing, does neither require stable fixation nor manual motor responses. For these reasons, we consider the present proof-of-concept sufficiently promising to verify its ability to serve as a screening tool for visual field defects in clinical practice.

Acknowledgments

This project has received funding from the Graduate School of Medical Sciences, Rijksuniversiteit Groningen and the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 641805 to FWC. RSS received a stipend from the Graduate School of Medical Sciences Abel Tasman program to work on the project. We are thankful to Nomdo Jansonius for his advice and help.